# AN EMPIRICAL MODELING APPROACH TO RECIDIVISM CLASSIFICATION

EDWARD DOW
CHARLES JONES
JACK MOTT
*Triant Psychometrics Corporation*

A nonlinear, exemplar-based empirical modeling methodology was applied to the problem of classifying relative levels of recidivism risk in a population of released offenders from the Wisconsin Department of Corrections. Issues related to extracting relatively pure classes of exemplars from relatively ambiguous data are detailed. Risk was defined as the associative match to one of two exemplar groups; higher or lower risk offenders. The area under the Receiver Operating Characteristic (ROC) curve for 620 offenders examined in the initial subgroup was .94. Comparable results were found with a smaller validation sample of 408 offenders known to be higher risk. Implications of controlling for risk factor patterns are discussed.

***Keywords:*** empirical modeling; recidivism; classification; high and low risk offenders; risk factors

Solomon and Camp (1993) listed a set of challenges that offender classification systems must effectively address: assist in the management of crowded prisons at the administrative level, augment decision-making quality, improve management and operations at the prison level, optimize the use of limited resources across the system, systematize the collection and analysis of data both for planning of future usage requirements and for security determination purposes, and provide a means to increase the assurance of public safety and protection. Researchers have continued to work on these issues. One area has been the progressive development of risk assessment tools.

A representative list of pioneering and current risk assessment tools include: the Client Case Staff Deployment (CCSD; Baird, Henke, &

CRIMINAL JUSTICE AND BEHAVIOR, Vol. 32 No. 2, April 2005 223-247
DOI: 10.1177/0093854804272892

Bemus, 1974); the Client Management Classification Report (CMC; Baird, Heinz, & Bemus, 1979); the Dangerous Behavior Rating Scheme (Webster & Menzies, 1993); the Historical, Clinical, and Risk Management violence risk assessment scheme (HCR-20; Webster & Polvi, 1995); the Lifestyle Criminality Screening Form (LCSF; Walters, White, & Denney, 1991); the Level of Service Inventory (LSI and LSI-R; Andrews, 1982), the Minnesota Sex Offender Screening Tool—Revised (Epperson, Kaul, & Hesselton, 1998); the Psychopathy Checklist (PCL and PCL-R; Hare, 1991); Rapid Risk Assessment for Sex Offence Recidivism (RRASOR; Hanson, 1997); the Spousal Assault Risk Assessment Guide (Kropp, Hart, Webster, & Eaves, 1994); the Static 99 (Hanson & Thornton, 1999); the Violent Offender Treatment Program Risk Assessment Scale (Ward & Dockerill, 1999); and the Violence Risk Appraisal Guide (VRAG; Harris, Rice, & Quinsey, 1993).

These instruments emphasize the use of risk factors that anticipate an offender's future behavior. However, the field is reaching a point where the range of accumulated risk factors is not likely to change greatly. Although nuanced observations about the drivers for recidivism will continue to emerge, some essential risk factors have been established. This assertion is evidenced by the overlap of risk factors in many of the instruments in the forefront of classification research (Barbaree, Seto, Langton, & Peacock, 2001; Kroner & Mills, 2001). For instance, the CCSD, CMC and the LSI-R make use of largely equivalent constructs for general recidivism. Similarly, the HCR-20, VRAG, and the Sex Offender Risk Appraisal Guide (SORAG) all include the PCL-R in deriving final scores. Kroner and Mills (2001) found a lack of statistical significance in predictive accuracy between PCL-R, LSI-R, HCR-20, VRAG, and the LCSF, which they attributed to content overlap.

## AN ALTERNATIVE APPROACH TO RISK CLASSIFICATION

Borum (1996) proffered recommendations to improve the clinical practice of risk assessment that included the improvement of assessment technology. His inclusion of risk assessment technology underscores the contention that important advances in classification accuracy arising from different ways of mathematically combining risk factors are essential.

This project introduces an approach to recidivism classification based on an advanced multivariate modeling algorithm currently used for applications such as fighter aircraft (Billings, 1990), nuclear reactor functioning (Mott & Blanch, 1992), fault detection (Mott, King, Monson, Olson, & Staffon, 1992), fault tolerance (Singer, King, & Mott, 1989a), real-time liquid metal reactor control (King & Mott, 1990), pattern recognition (King, Radtke, & Mott, 1988; Mott & Young, 1987; Mott, Young, & King, 1987; Singer, King, & Mott, 1989b), and general system's state analysis (King et al., 1988; Mott, King, & Radtke, 1988).

Dow (1995) evaluated the applicability of the above modeling concepts to non-machine environments by comparing a commercially available version of the underlying modeling technology referenced above (Teranet, 1992) to discriminant analysis. In that study, data derived by a consensus of four experts was used to classify the recommendations for academic "tracking" of students. The modeling algorithm outperformed traditional discriminant analysis in classification accuracy by more than 50%.

### A FEASIBILITY STUDY

A cornerstone of the Wisconsin Department of Corrections' (WIDOC) approach to the risk assessment is the Client Management Classification (CMC) Report (Lauen, 1997) and its progenitor, the Case Classification Staff Deployment (CCSD) project (Baird et al., 1974). The purpose of this study was to determine whether an empirical modeling approach could use offenders' CCSD data to determine which offenders posed a lower or higher risk for committing new crimes upon release.

## METHOD

### PARTICIPANTS

Feasibility data was collected at WIDOC. The initial data encompassed 107,041 admission records and 29,342 CCSD records. These

**TABLE 1: Recidivism Over Time**

| *Years Between* | *Instances of Parole Revocation (cumulative %)*[a] | *New Convictions (cumulative %)*[b] |
|---|---|---|
| 1 | 48.64 | 44.78 |
| 2 | 70.85 | 67.08 |
| 3 | 81.07 | 78.08 |
| 4 | 87.10 | 84.98 |
| 5 | 90.77 | 89.24 |
| 6 | 93.27 | 92.06 |
| 7 | 94.95 | 94.02 |
| 8 | 96.24 | 95.57 |
| 9 | 97.22 | 96.78 |
| 10 | 97.90 | 97.64 |

*Note*. Some offenders took longer than 10 years to be readmitted.
a. $n$ = 41,322.
b. $n$ = 22,369.

records were for all admissions dating back approximately from November 1969 to October 2001. This population was comprised of 43,176 offenders who had only one admission and 22,544 recidivists with between 2 and 13 admissions, with a mean of 2.86 ($SD$ = 1.40) admissions per recidivist. Of the recidivist population, 9,668 were reconvicted with between 2 and 13 admissions, between 2 and 9 new sentences, and a mean of 3.31 ($SD$ = 1.70) admissions per reconvicted recidivist. The majority of the CCSD records were for probationers, but 7,125 were for parolees.

In the service of tracking the return rate of offenders back to incarceration, it seems useful to present the trends as they unfolded in this data set. A total of 65,720 offenders were examined. Table 1 lists the cumulative instances, by release duration, of post-release return to prison. As can be seen from Table 1, by year 3, a bit more than three quarters of those who would be reconvicted for new crimes had been identified, leaving one quarter in a latency period before identification.

This project focused on parolees and their CCSD reports. Of this collection, 5,941 cases contained all the relevant information necessary for modeling. The ages of these 5,941 offenders ranged from 16 to 89, with a mean of 33.13 ($SD$ = 8.50). Time served ranged from 0 to

27-plus years, with a mean of 4.66 (*SD* = 3.7); the number of sentences from 1 to 8, with the mean number of sentences 1.10 (*SD* = .96); and the number of admissions ranged from 1 to 11, with the mean number of admissions 2.39 (*SD* = 1.70). The sample included 5,357 males and 584 females.

### PROCEDURE

Each offender was linked to two pieces of data: A CCSD and a crime report contained in the admission records. A simple merge of the Wisconsin CCSD data and release records was performed and denoted as the Offender Data Array (ODA). Appendix A lists the variables used in this study.

Many risk assessment tools consolidate crimes into major or minor offenses (e.g., felony or misdemeanor). This type of consolidation loses a great deal of information. For instance, it tends to weight crimes like robbery equal to manslaughter, as most jurisdictions would consider both to be felonies. Another method relies on the criminal code. This also has inherent problems stemming from charging practices that often result in the imposition of a lesser charge on an offender. A consolidated crime index of the penal code was constructed to address these issues (see Appendix B).

### DISCRIMINANT FUNCTION: EMPIRICAL MODELING REQUIREMENTS

Classifying offenders requires that an offender population possess a criterial characteristic of interest (e.g., recidivism). Identifying this target distinction introduces the possibility of developing a discriminant function. It is designed in such a fashion that its resulting value correlates strongly with both the presence and the absence of that particular characteristic. Success is then defined by a contingency: If you know the level of the variable, you will also know the relative presence or absence of the known characteristic. This study searched for a discriminant function to distinguish higher- and lower-risk groups from each other.

**DIFFERENCES:**
**EMPIRICAL MODELING AND**
**CLASSIC STATISTICAL MODELING**

Three primary features distinguish the methodology applied in this study from other traditional approaches.

*Data patterns:* A typical risk assessment procedure entails formulating a scale, scoring the items, adding up items, and making inferences based on the resulting index score. In contrast, this process uses patterns in risk factor data and links those patterns together into dynamic classification models.

*Exemplars:* Exemplars are ideal examples of a class. This methodology both refines and uses the best available examples of higher- and lower-risk groups to form an empirically based reference library (REFLIB) used for dynamic modeling. In contrast, traditional approaches create static norm tables.

*Models:* Models are constructed from patterns of data associated with the criterial groups. Briefly, a subset of exemplars is selected whose data patterns most closely matched the data patterns of an offender being modeled. The degree of match is then used to calculate risk-level classifications. In concept, this is a nearest-neighbor approach, but unlike classical approaches using Euclidean distance to select nearest neighbors, this method utilized a nonlinear distance measure that optimizes the similarity in patterns rather than minimizing differences. This precluded any single variable from dominating the discriminant function.

Putting these three features together in a more detailed fashion, the modeling procedure operated as follows: When a specified candidate is selected for the modeling process, a small number of comparable offenders are selected from the exemplars comprising the REFLIB. The comparable offenders are the exemplars having data record patterns most like that of the specified candidate. Next, a series of optimizing calculations are performed. These calculations determine the weighting for each of the selected exemplar data record patterns and variables. These patterns are combined to produce a modeled data pattern most like the candidate's actual data record pattern.

An adage from psychology reads: If you want to anticipate what someone is likely do in the future, examine what that person did, under similar circumstances, in the past. Unfortunately, this adage is not helpful in its current form because the behavioral sample generated by an offender concerning release is limited. Consequently, the adage was revised to read: If you want to anticipate what someone is likely do in the future, examine what a group of other comparable people did, under similar circumstances, in the past. The utility of the revised adage depends on the relationship of the comparable offenders' risk pattern to the outcome of interest.

#### CREATION OF AN EXEMPLAR LIBRARY UTILIZED IN MODELING

Classifications based on examples that carry a lot of extraneous information will yield confounded or even incorrect results. However, if good examples could be extracted, better results would be expected. This approach was used to generate the exemplars needed for the present exemplar-based modeling. In effect, this process started with a pool of examples—typical illustrative instances of higher- and lower-risk cases muddled together with cases that fall into an unclear class designation. Because the modeling process used in this study can be used as a noise filtration algorithm, it was used to distill relatively ideal representative exemplars of the pure classes to facilitate modeling. Thus, a move was made from a reference population of examples to a REFLIB of exemplars that provided the basis for modeling. The REFLIB had to contain exemplars indicative of individuals for whom we would be expected to generate models. In the current case of recidivism, this meant populating the library with two groups: offenders who did and who did not recommit crimes after they were released.

#### AN INEVITABLE DATA PROBLEM: WHAT CAN WE KNOW ABOUT OUR EXAMPLES?

The ODA represented some relatively knowable information about each offender: for example, the crime committed, the age of the offender, the number of prior supervised releases. Everything in the CCSD revealed a direct observation or an inference closely tied to a direct observation. It was knowable to the extent that someone supplying the information could reliably observe or acquire and then document the information. This was not so true of the second set of data. The second data set provided information about what occurred after release as determined by dispositional records, such as arrest and prison admission records. In the data received, post-release history was classified into two categories: reincarcerated and not reincarcerated. If a released offender had been readmitted, two additional subcategories were designated: readmission because of a new conviction or readmission because of a parole revocation.

There was an inherent problem with this linkage. As indicated, the first set of ODA data was based on largely observable information; we could know what each designation indicated. In contrast, this second set of data (post-release history) was the result of a complex dispositional decision-making process. At its foundation, this process required that an offender be caught for a crime and then successfully prosecuted, plea-bargained, and so forth. Herein lies the problem for selecting exemplars. Classifying by disposition is not necessarily an accurate summary of an offender's behavior. For example, not being readmitted to prison subsequent to release does not preclude the possibility that criminal activity had occurred. Some offenders do not get caught and some do not get prosecuted. This presented a verification problem. Exemplar-based modeling uses historical data patterns and their associated outcomes to model new patterns for which the outcomes are unknown. If the historical patterns have unclear outcomes, a source of uncertainty will surface.

## THE IMPACT OF TIME AS A COMPLICATING FACTOR ON DEFINING GROUPS

Establishing membership in one of the three groups was, to a nontrivial extent, a function of time. In the full data set, there were offenders who returned to prison the day after release. However, most required months or years to fully establish their membership in one of the three groups. This highlights the danger of classifying recidivism potential for offenders based on short at-large periods (periods of less than 2 years will exclude approximately 33% of the offenders that will ultimately go on to commit a new crime).

### USING WHAT WE HAVE TO MODEL WHAT WE CAN

Of our sample of 5,941, 3,267 fell into the no-readmission group, 1,718 into the readmitted-without-a-new-conviction group, and 956 into the readmitted-with-a-new-conviction group.

### NOT ALL EXAMPLES ARE GOOD EXEMPLARS: AN APPROACH TO REFINING

As defined by the purpose of this study, a collection of exemplars of offenders who did and who did not commit a new crime was sought. Unfortunately, the data in the reference group did not directly designate those classes because not all offenders who had committed crimes were captured, nor can it be conclusively assumed that no innocent people were improperly convicted. To designate the classes, exemplar-based modeling was used to generate a partially refined transitional exemplar library.

The revocation group was the least useful of the three groups because it was not clear what its membership exemplified. Some of its members had committed new crimes, but were not charged, yet returned for technical violations. Others had not committed crimes but were revoked nonetheless for technical violations. In contrast, the not-readmitted group and the new-conviction group offered better examples of the two classes needed to start the modeling process because the not-readmitted group stayed in the community for several years without any detected incidents, whereas the new-conviction group had been detected, arrested, successfully prosecuted, and sentenced, thereby insuring some degree of confidence of their previous behavior.

It was expected that the new-conviction examples would provide a good source of exemplars of offenders who would recommit crimes. After all, the members of this class had been charged, prosecuted, and convicted of new crimes. This group, labeled higher-risk transitional exemplars, comprised 816 offenders. With respect to the not-readmitted group, some were not out of prison long enough to reveal their true colors. Experience has shown that it takes about 5 years for the identification of these three groups to emerge (see Table 1). Using this time trait (i.e., observations were truncated after 5 years) as an additional selection factor, an examination of the pool of non-readmitted examples was conducted. Ideally, these exemplars, labeled lower-risk transitional exemplars, would have been out at least 2 years. Of the 3,267 total, there were only 646 who fit the criterion for transitional exemplars for this group.

#### A TRANSITIONAL DEFINITION OF RISK

The above two groups of transitional exemplars (i.e., high-risk transitional new-conviction and low-risk transitional not-readmitted) were combined to be used as a transitional REFLIB. Although not the final REFLIB, using them during an interim methodological step provided a modest increase in precision. It enabled the filtering of the data for the distinctive informational patterns that distinguish the final REFLIB of good exemplars from mere examples.

For this interim methodological step, an initial yardstick was required for differentiating the potential risk of the offenders in these two selected groups. The number of admissions to prison seemed to provide the best proxy for degree of risk. The range of admissions to prison in the refined groups ranged from 0 to 11 readmissions, with a mean of 3.97 (*SD* = 3.98).

#### CLASSIFYING GROUP MEMBERSHIP: COMPARISON TO COMPARABLE OFFENDERS

The constructed transitional REFLIB, comprised of a new-conviction group and the not-readmitted group, was used in the next step to generate a modeled group membership. Offenders with a new conviction provided the cleanest information. With that in mind, a validation sample of 408 (dividing the high-risk transitional exemplar group of 816 in half) was generated as a way to provide a validation sample for later use. The membership of the two transitional exemplar groups is summarized in Table 2.

**TABLE 2: Transitional Exemplars Used for the Interim Methodological Step**

| *Characteristics* | *Sample* n | *Yield* |
|---|---|---|
| Higher-risk transitional exemplars from the new-conviction group: Offenders released between January 1999 and October 2000, who have more than one new conviction. | 408 | Selected to provide a concentrated—fairly pure—collection of higher-risk offenders. |
| Lower-risk transitional exemplars from the not-readmitted group: Offenders released between February 1980 and January 1999 who have no readmissions for either conviction or revocation. | 646 | Selected because it will have at least some lower-risk offenders who cannot otherwise be identified. |

### MODELING OFFENDERS IN THE NEW-CONVICTION GROUP

Exemplar-based modeling was used to generate a model of each of the 408 higher-risk candidates. Ten comparable offenders were selected from the transitional exemplar library by this process to model number of admission. No self-modeling was allowed (i.e., an offender model was not affected by having his or her data influence the construction of the discriminant function). The results are shown in Figure 1.

This group had only higher-risk members, so that the actual number of admissions was always greater than 1. As expected, the modeled number of admissions for members of this group is almost always greater than 1.

### MODELING THE OFFENDERS NOT IN THE NEW-CONVICTION GROUP

Similarly, the same modeling technique described above was used to generate a model of each of the 646 lower-risk candidates. The results are presented in Figure 2. Please note that Figure 1 and 2 have different vertical scales.

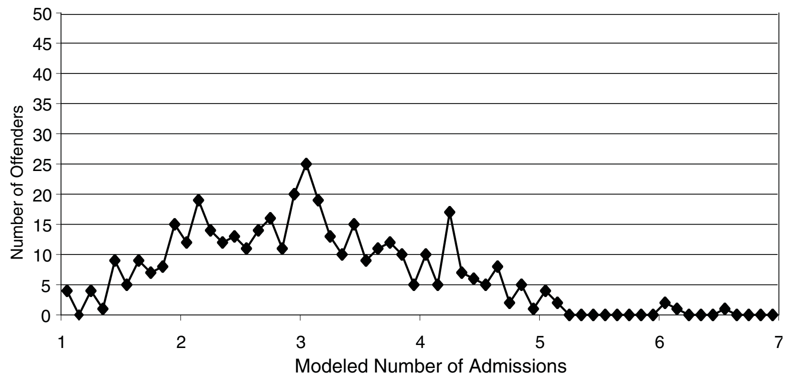


**Figure 1: Models of Higher-Risk Offender Candidates (408 Offenders Released Between January 1999 and October 2001 Who Were Reconvicted by October 2001)**

This group was less pure than the first group. Its members had not been reconvicted, but it was expected that some of them would commit—or had already committed—new crimes. Thus, the actual number of admissions was unreliable, and we expected the modeled number of admissions to spread out. Notice, though, that this modeling step produced a significant number of offenders with a modeled number of admissions very nearly equal to 1 (the left-hand spike).

#### PATTERNS IN THE DISTRIBUTIONS

Figures 1 and 2 provided a visual examination of the modeling results for lower-risk and higher-risk offenders. There were discernible trends. The model of higher-risk transitional exemplars (Figure 1) did not peak and was not biased toward the left. The model of lower-risk transitional exemplars (Figure 2) was biased and peaked toward the left.

These two distributions allowed for further decisions to be made in refining the selection of exemplars for the final REFLIB. Given the fact that this was a feasibility study, a simple differentiation, a cutpoint of 1.15, was selected. Thus, for those higher-risk offenders modeled in

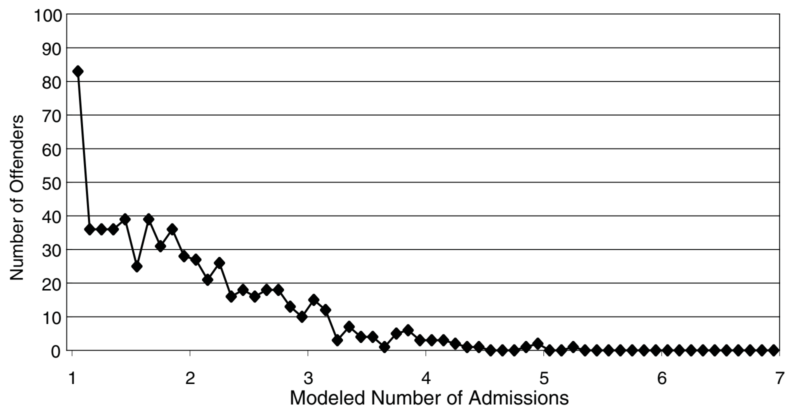


**Figure 2: Models of Lower-Risk Offender Candidates (646 Offenders Released Between February 1980 and January 1999 Who Had Not Been Readmitted as of October 2001)**

Figure 1, transitional exemplars with modeled values below 1.15 were removed as they modeled like lower-risk offenders. The procedure left 390 final higher-risk exemplars from the original 408 transitional exemplars. In contrast, the lower-risk offenders modeled in Figure 2, transitional exemplars with modeled values above 1.15 were removed as they modeled like higher-risk offenders. This procedure left 230 final lower-risk exemplars from the original 646 transitional exemplars.

The removal of 416 offenders was generally expected given the uncertainty inherent in the source of exemplars. This interim methodological step enabled the removal of salient features more consistent with higher-risk exemplar patterns from the lower-risk group of exemplars. In effect, identification of the lower-risk group was made by not having characteristics salient to the higher-risk group.

### TRANSFORMING LOWER-RISK AND HIGHER-RISK MODELING INTO RISK VALUES

Through the transitional modeling process, cleaner exemplars were selected for the permanent REFLIB. This enabled the final REFLIB to be produced by pooling both exemplar groups, assigning the lower-risk exemplars a value of 1 and the higher-risk exemplars a value of 2. Thus, instead of using the data pattern library to model the number of admissions, the data pattern library was used to model risk class. This generated a continuum of risk ranging from 1.00 (indicating a greater association with the lower-risk exemplar group) to 2.00 (indicating a greater association with the higher-risk exemplar group).

## RESULTS

### MODELING LOWER-RISK OFFENDERS

Using the combined REFLIB, 230 lower-risk exemplars were modeled and plotted producing the resulting modeled-risk values as shown in Figure 3.

### MODELING HIGHER-RISK OFFENDERS

Using our combined REFLIB, 390 higher-risk exemplars were modeled and plotted, producing the resulting modeled-risk values as shown in Figure 4.

### SOME IMPORTANT FEATURES OF THE LOWER-RISK AND HIGHER-RISK FIGURES

The curve in Figure 3 can be thought of as a confidence level that an offender with a value less than the modeled-risk level is associated with the class of lower-risk offenders. Similarly, the curve in Figure 4 provides a confidence level that an offender with a value higher than the modeled risk is associated with the class of higher-risk offenders.

A modeled-risk value of about 1.55 (where these two curves cross one another when plotted together on the same graph) can be thought of as a separator point. Offenders with modeled-risk scores below that

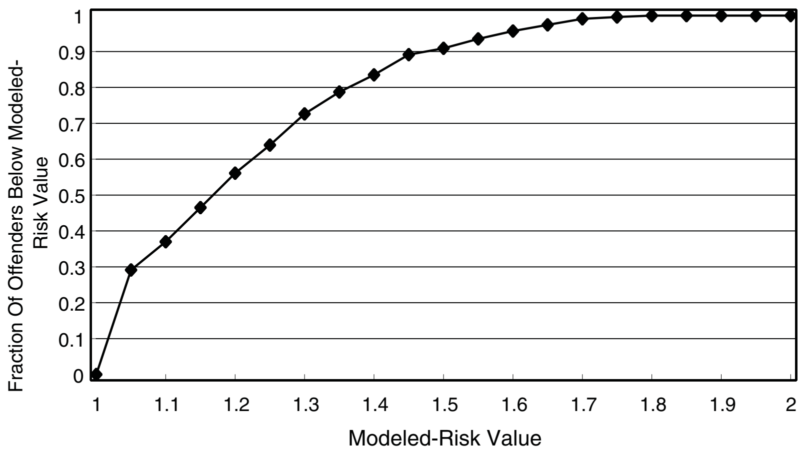


**Figure 3: Models of Lower-Risk Offender Exemplars (230 Offenders Released Between February 1980 and January 1999 Who Had Not Been Readmitted as of October 2001 and Who Contain None of the Salient Features of Higher-Risk Offenders, Their Actual-Risk Value = 1)**

level matched patterns consistent with the lower-risk exemplars at an approximate confidence level of 93%. Similarly, offenders with modeled-risk scores above that level matched patterns consistent with the higher-risk exemplars at an approximate confidence level of 93%.

### DEFINING RISK MORE SPECIFICALLY

Putting all of this a bit differently, we can more clearly define some important concepts. Modeled-risk values provided a statistical means to determine the degree of match between a release candidate, with unknown recidivism prospects, and either of two risk level groups with known outcome histories.

The lower-risk group consisted of exemplar offenders, as best could be determined from records covering at least a 2-year release period. These offenders were not returned to prison for technical violations, new convictions, nor did they have the salient features of the higher-risk group. The higher-risk group consisted of exemplar of-

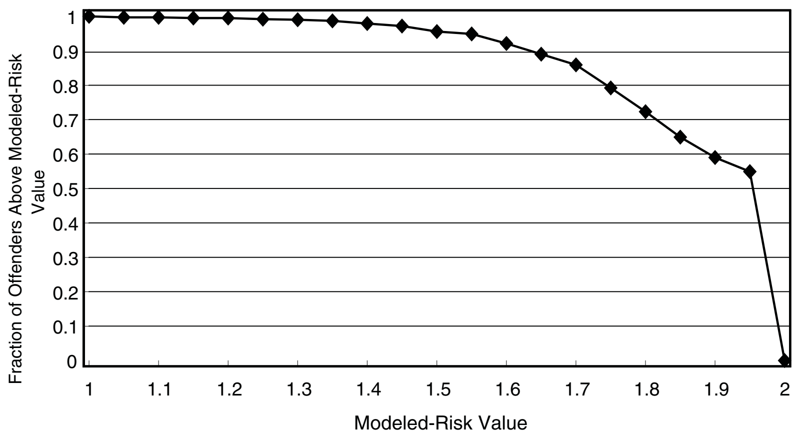


**Figure 4: Models of Higher-Risk Offender Exemplars (390 Offenders Released Between January 1999 and October 2001 Who Were Reconvicted by October 2001 and Who Contain None of the Salient Features of Lower-Risk Offenders, Their Actual-Risk Value = 2)**

fenders who had returned to prison within 2 years with at least one new conviction without the salient features of the lower-risk group.

The modeled-risk value located an offender, relative to all other modeled offenders, along a continuum from 1.00 (lower-risk) to 2.00 (higher-risk). These values represented the relative fit between the candidate and the unique distribution of modeled-risk values characteristic of each group of exemplars.

### A RECEIVER OPERATING CHARACTERISTIC (ROC) ANALYSIS

Used since the 1940s for electronic signal detection systems, the Receiver Operating Characteristic (ROC) table is the basis for a standard statistical procedure used for determining the effectiveness of a classification (see Swets, Dawes, & Monahan, 2000, for an excellent ROC primer). The ROC table presented here indicates the match between actual and modeled values for both the lower-risk group and the higher-risk group.

The fraction of offenders below a modeled-risk value in Figure 3 is also the true-positive ratio derived from a ROC table. The fraction of offenders above a modeled-risk value in Figure 4 is also the complement of the false-positive ratio derived from an ROC table.

The actual risk was designated as 1 for the 230 lower-risk exemplars and 2 for the 390 higher-risk exemplars discussed above. If an offender had a modeled-risk value that was less than a threshold, then he or she was classified as a lower-risk offender. If an offender had a modeled-risk value that was greater than a threshold, then he or she was classified as a higher-risk offender. As the threshold changes, the number of offenders classified as lower-risk or higher-risk changes, and correspondingly, the numbers in each section of the ROC table change.

The ROC curve is the true-positive ratio as a function of the false-positive ratio. It provides a graphic representation of how effectively the risk groups were classified. For this application, the ROC curve is shown in Figure 5 and is equal to .94. An Area Under the Curve (AUC) equal to .5 would indicate that classification as low-risk or high-risk was purely random, whereas an AUC equal to 1 would indicate perfect classification.

This ROC (AUC = .94) is a representation of how well the development sample of lower-risk exemplars and higher-risk exemplars were separated while using leave-one-out modeling.

An ROC curve for the 408 offenders in the validation sample could not be generated because only a pool of higher-risk candidates was available for modeling. As argued, a clean set of known lower-risk offenders is difficult to establish, hence, hard to generate for comparison sake in this feasibility study. As a result, all of the lower-risk offenders were used to generate our final REFLIB. However, some results are presented from the modeling process in the form of a graph that is comparable to Figure 3, the model of the 408 higher-risk exemplars. The graph is presented in Figure 6.

Comparing the curve graphed in Figure 6 (the validation sample of higher-risk offenders) to that in Figure 4 (the group of higher-risk offenders used as final exemplars) reveals a small degradation in the modeled results. Some comparisons between the two, at different cutpoints, are presented in Table 3.

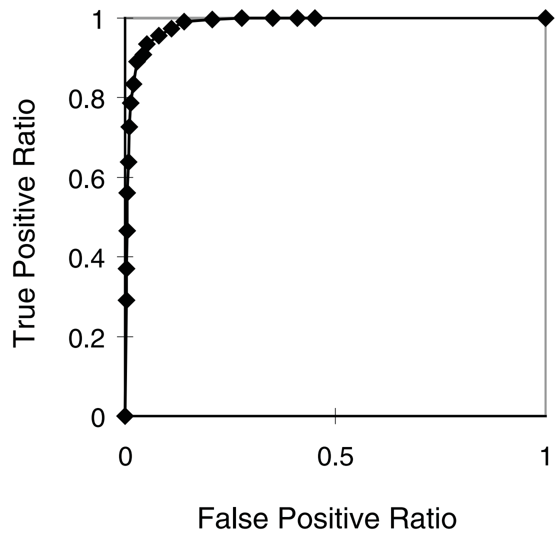


**Figure 5: Receiver Operating Characteristic (ROC) Curve for Lower-Risk and Higher-Risk Offender Exemplars (Area Under Curve = .94)**

## DISCUSSION

Clements (1996) described correctional classification as the process of subdividing offenders into useful groups. Furthermore, he stated, "Somewhere between the extremes of all offenders are alike and each offender is unique lies a system (or systems) of categorization along pertinent dimensions that will prove of value in reaching correctional goals" (p. 123). Clements rightly observed that current risk assessment tools do not generate highly individualized profiles or take advantage of unique combinations of risk factors. This study demonstrated the application of an empirical modeling approach to the task of risk classification. This approach differentiated offenders who were more or less prone to commit new crimes by controlling for the underlying risk factor pattern presented by each of the offenders.

Of significance was the fact that the risk factors driving this modeling approach were a modified set of the Wisconsin Risk/Needs and CMC. As recent as February 2001, these scales were by far the dom-

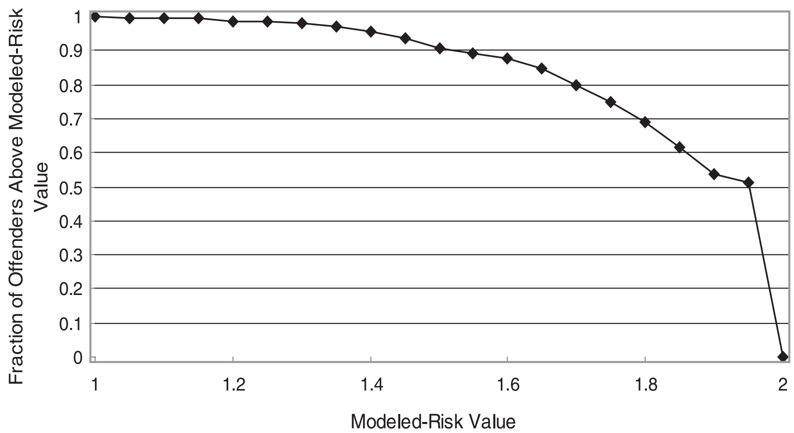


**Figure 5: Models of the Higher-Risk Offender Validation Set (408 Offenders Released Between January 1999 and October 2001 Who Were Reconvicted by October 2001, This Group Contains None of the Higher-Risk Offender Exemplars)**

**TABLE 3: Comparing the Model Results of Two Higher-Risk Groups**

| | *Cutpoint* | *Number Modeled* | *Correct Classification* |
|---|---|---|---|
| Exemplars | 1.45 | 379 | 92.9% |
| | 1.50 | 373 | 91.4% |
| | 1.55 | 370 | 90.7% |
| Validation | 1.45 | 381 | 93.4% |
| | 1.50 | 370 | 90.7% |
| | 1.55 | 363 | 89.0% |

*Note.* $n$ = 816: 408 exemplars and 408 validation-sample members.

inant frameworks used by probation/parole agencies in the United States (Hubbard, Travis, & Latessa, 2001). Because many jurisdictions collect the core set of risk factors used in this study, this classification approach may have wide application.

As presented, the core issue associated with risk assessment has not been a lack of consensus about risk factors associated with recidivism but rather, the reliance on adding up risk factors to generate an index score. This reliance has had the undesirable effect of assigning fixed weights to risk factors regardless of the resiliency of an offender. An offender's life experience, whether male or female, will modulate how an offender is affected by any particular constellation of risk factors. In many ways, an offender's risk constellation represents his or her range of resiliency. As such, assessment tools that assign weights to risk factors without taking into account the overall risk pattern exhibited by the offender should be suspect.

It is not difficult to imagine two different offenders having the same score on an instrument like the LSI-R but getting that score via different risk factors. Knowing that risk factors correlate differently with recidivism, does it make sense to weight those factors with the same point value? Does it follow that the same recidivism rate applies simply because two offenders have the same score? The failing of this assumption is belied by sex-based norms as men and women having the same index score often represent different levels of risk.

Fowler (1993) implicitly understood this issue in the context of classifying female offenders. Female offenders, who are subjected to classification tools developed on male offenders, are often overclassified. The reason for the overclassification can be traced directly to the failure of index scores and the false assumption that if two offenders have the same score, then they present the same risk. This same reasoning can be extended to apply within the same sex as well; not all males having the same index score necessarily present the same risk.

Although budgetary concerns now facing corrections will be resolved, they do draw attention to the need to better the business of corrections. Perhaps nowhere is this issue more salient than in the assessment of risk. The overclassification of offenders has direct financial consequences, whereas underclassification presents a threat to the community at large. Methods of controlling for the underlying pattern of risk factors allow classification systems to be applied to a wider range of offenders, regardless of sex, race, or reason for incarceration while increasing accuracy. The exemplar-based modeling approach used in this study maximizes the similarity in risk factor patterns. In addition, it creates an empirically based model of the recidivism associated with a particular pattern of risk factors that may be useful during supervision.

One final implication of the modeling method now described concerns dynamic and static risk variables. This modeling approach maximizes the similarity in an offender's pattern to a REFLIB, rather than minimizing differences. The subtle effects of a change in even one risk factor can potentially change the selection of the comparable offenders used to construct a model. As the composition of the selected comparable offenders changes, so can the risk values. In contrast, offenders assessed with a risk tool generating an index score must often have changes to several items to cross into a different risk group before detecting a change.

As with all research, improvements can be made. This study used data that had already been collected. Although the number of risk factors used for modeling was large, this study made no effort to identify a reduced set of risk factors. This decision was made because the underlying modeling technology automatically de-emphasizes irrelevant data. Consequently, this had little impact on demonstrating the utility of this approach. Additional research is presently underway to determine whether exemplar-based modeling can effectively identify offenders presenting the greatest potential of risk reduction when given treatment. Future research that examines the interactions among the variable set and the effects of reducing that variable set is warranted.

In conclusion, the modeling methodology used in this study addresses the notion that all offenders are "alike," while also addressing the notion that offenders have "unique" qualities. The ability to create optimal dynamic models may provide the next step in risk assessment technology, further bridging the gap in the ability to accurately identify offenders more prone to recidivate than others.

## APPENDIX A
## Variables Collected to Drive the Modeling Process

| No. | Variable Description | 41 | Will client be referred to Clinical Services |
|---|---|---|---|
| 1 | Academic/Vocational Skills | 42 | Primary Client Management Classification |
| 2 | Employment | 43 | Living Arrangement |
| 3 | Financial Management | 44 | Number of Dependents |
| 4 | Marital/Family Relationships | 45 | Making Support Payments |
| 5 | Companions | 46 | Need Child Care |
| 6 | Emotional Stability | 47 | Veteran |
| 7 | Alcohol Usage | 48 | Amount of Time Employed |
| 8 | Other Drug Involvement | 49 | Months at Current Job |
| 9 | Mental Ability | 50 | Job Classification |
| 10 | Health | 51 | Current Gross Monthly Income (Wages) |
| 11 | Sexual Behavior | 52 | Job Training Wanted by Client |
| 12 | Agent's Impression of Client's Needs | 53 | Last Grade Completed |
| 13 | Num of Address Changes in Last 12 Months | 54 | Num Prior Misdr Convictions (Adult) |
| 14 | Percentage Time Employed in Last 12 Months | 55 | Num Previous Misdr Probations (Adult) |
| 15 | Alcohol Usage Problems Prior to incarceration for parolees | 56 | Num Previous Felony Probations (Adult) |
| 16 | Other Drug Usage Problems Prior to incarceration for parolees | 57 | Num Times Previously Released on Parole |
| 17 | Attitude | 58 | Num Prior Incarcerations for One Year or Longer in a Federal prison |
| 18 | Age at First Conviction or Juvenile Adjudications | 59 | Disabled Aid/Worker's Comp |
| 19 | Num of Prior Periods of Probation/Parole | 60 | Social Security (SSI) |
| 20 | Num of Prior Probation/Parole Revocations | 61 | VA Benefits |
| 21 | Num of Prior Felony Convictions | 62 | Unemployment Comp |
| 22 | Convictions or Juvenile Adjudications for Include current offense | 63 | AFDC |
| 23 | Convictions or Juvenile Adjudication for Assaultive | 64 | General Relief |
| 24 | Self-concept problems -low self-esteem | 65 | Other |
| 25 | Self-concept problems -grandiosity | 66 | Type of Admission |
| 26 | Interpersonal problems with -peers | 67 | Crime Severity 1a offense 1 |
| 27 | Interpersonal problems with -authority | 68 | Crime Severity 1b offense 1 |
| 28 | Interpersonal problems with -family | 69 | Crime Severity 1c offense 1 |
| 29 | Emotional problems -depression | 70 | Crime Severity 1d offense 1 |
| 30 | Emotional problems -history of psychotic | 71 | Crime Severity 1a offense 2 |
| 31 | Emotional problems -anxiety | 72 | Crime Severity 1b offense 2 |
| 32 | Mental Health Treatment History -inpatient | 73 | Crime Severity 1c offense 2 |
| 33 | Mental Health Treatment History -outpatient | 74 | Crime Severity 1d offense 2 |
| 34 | Destructive behavior -self | 76 | Institution Security Level |
| 35 | Destructive behavior -property | 77 | Male/Female |
| 36 | Destructive behavior -persons / assaultive | 78 | Age |
| 37 | Unusual behavior / thought disorder | 79 | Length of Incarceration |
| 38 | Learning disability / mental retardation | | |
| 39 | Criminal / antisocial value system | | |
| 40 | Other (if referring, specify in box) | | |

## APPENDIX B
## Crime Severity Index (CSI) Used to Consolidate Legal Code

| Code | **CSI1** Type |
|---|---|
| 0 | Not selected |
| 1 | Homicide with intent |
| 2 | Homicide without intent |
| 3 | Sex assault adult |
| 4 | Sex assault child |
| 5 | Sex crime other (prostitution, child porn) |
| 6 | Kidnapping |
| 7 | Assault |
| 8 | Robbery |
| 9 | Other Arson |
| 10 | Theft vehicle |
| 11 | Theft non-vehicle |
| 12 | Fraud |
| 13 | Drug possession |
| 14 | Drug trafficking |
| 15 | Weapon |
| 16 | Arson |
| 17 | Other violent |
| 18 | Other nonviolent |
| 19 | Escape |
| 20 | Neglect of child |

| Code | **CSI2** Macro Type |
|---|---|
| 1 | Violent |
| 2 | Property |
| 3 | Drug |
| 4 | Weapons |
| 5 | Escape |
| 6 | Other |

| Code | **CSI3** Weapon |
|---|---|
| 0 | Not Present |
| 1 | Present |

| Code | **CSI4** Intent |
|---|---|
| 0 | Not Present |
| 1 | Present |

## REFERENCES


Andrews, D. A. (1982). *The Level of Supervision Inventory (LSI): The first follow-up*. Ottawa, Canada: Carleton University, Department of Psychology.

Barbaree, H. E., Seto, M. C., Langton, C. M., & Peacock, E. J. (2001). Evaluating the predictive accuracy of six risk assessment instruments for adult sex offenders. *Criminal Justice and Behavior*, *28*, 490-521.

Baird, S. C., Heinz, R. C., & Bemus, B. J. (1979). *The Wisconsin Case Classification/Staff Deployment Project: A two-year follow-up report*. Madison, WI: Bureau of Community Corrections.

Baird, S. C., Henke, T., & Bemus, B. J. (1974). *The Wisconsin workload deployment project: Final report*. Madison, WI: Bureau of Community Corrections.

Billings, V. (Producer). (1990). *Northrop/McDonnell Douglas: ATF-23 Project—Unclassified video* [Videotape]. (Available from Triant Technologies Inc., 20 Townsite Road, 2nd floor, Nanaimo, BC, Canada, V9S 5T7)

Borum, R. (1996). Improving the clinical practice of violence: Risk assessment, technology, guidelines, and training. *American Psychologist*, *51*, 945-956.

Clements, C. B. (1996). Offender classification: Two decades of progress. *Criminal Justice and Behavior*, *23*, 121-143.

Dow, E. A. (1995). *A comparative evaluation of discriminant analysis and the universal process model in counseling psychology*. Unpublished doctoral dissertation, University of Wisconsin, Madison.

Epperson, D. L., Kaul, J. D., & Hesselton, D. (1998). *Final report of the development of the Minnesota Sex Offender Screening Tool—Revised (MnSOST-R)*. Paper presented at the 17th Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, Vancouver, British Columbia, Canada.

Fowler, L. T. (1993). What classification for women? In American Correctional Association (Ed.), *Classification: A tool for managing today's offenders* (pp. 37-45). Arlington, VA: American Correctional Association.

Hanson, R. K. (1997). *The development of a brief actuarial risk scale for sexual offence recidivism*. Ottawa: Department of the Solicitor General of Canada.

Hanson, R. K., & Thornton, D. M. (1999). *Static 99: Improving actuarial risk assessments for sex offenders*. Ottawa: Public Works and Government Services Canada.

Hare, R. D. (1991). *The Hare Psychopathy Checklist—Revised*. Toronto, Canada: Multi-Health Systems.

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1993). Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. *Criminal Justice and Behavior*, *20*, 315-335.

Hubbard, D. J., Travis, L. F., & Latessa, E. J. (2001). *Case classification in community corrections: A national survey of the state of the art* (NIJ Grant 98-IJ-CX-0008). Washington, DC: National Institute of Justice.

King, R., & Mott, J. (1990, August). *Real-time LMR control parameter estimation using advanced adaptive synthesis*. Paper presented at the proceedings of the 1990 International Fast Reactor Safety Meeting, Snowbird, UT.

King, R., Radtke, W., & Mott, J. (1988, August). *Pattern recognition system application to plant life extension*. Paper presented at the American Nuclear Society Topical Meeting on Nuclear Power Plant Life Extension, Snowbird, UT.

Kroner, D. G., & Mills, J. F. (2001). The accuracy of five risk appraisal instruments in predicting institutional misconduct and new convictions. *Criminal Justice and Behavior*, *28*, 471-479.

Kropp, P. R., Hart, S. D., Webster, C. D., & Eaves, D. (1994). *Manual for the Spousal Assault Risk Assessment Guide*. Vancouver, Canada: The British Columbia Institute on Family Violence.

Lauen, R. J. (1997). *Positive approaches to corrections: Research, policy and practice*. Lanham, MD: American Correctional Association.

Mott, J., & Blanch, P. (1992, May). *Feedwater flow estimation via sample-based modeling*. Paper presented at the proceedings of the 8th Power Plant Dynamics, Control, and Testing Symposium, Knoxville, TN.

Mott, J., King, R., Monson, L., Olson, D., & Staffon, J. (1992, May). *A universal fault-tolerant, non-linear analytic network for modeling and fault detection*. Paper presented at the proceedings of the 8th Power Plant Dynamics, Control, and Testing Symposium, Knoxville, TN.

Mott, J., King, R., & Radtke, W. (1988, August). *A generalized system state analyzer for plant surveillance*. Paper presented at the proceedings of the American Nuclear Society/European Nuclear Society (ANS/ENS) International Topical meeting on Artificial Intelligence and other Innovative Computer Applications in the Nuclear Industry, Snowbird, UT.

Mott, J., & Young, R. (1987, September). *Pattern-recognition software for detecting the onset of failures in complex systems*. Paper presented at the proceedings of the 42nd Meeting of Mechanical Failures Prevention Group, National Bureau of Standards, Gaithersburg, MD.

Mott, J., Young, R., & King, R. (1987, August). *Pattern-recognition software for plant surveillance*. Paper presented at the proceedings of the ANS/ENS International Meeting on Nuclear Power Plant Operation, Chicago.

Singer, R., King, R., & Mott, J. (1989a, November). *An analytical approach to achieving fault tolerance in nuclear power plant sensors*. Paper presented at the American Nuclear Society Winter Meeting, San Francisco, CA.

Singer, R., King, R., & Mott, J. (1989b, September). *Use of pattern recognition scheme to compensate for critical sensor failures*. Paper presented at the proceedings of the First International Machinery Monitoring Diagnostic Conference, Las Vegas, NV.

Solomon, L., & Camp, A. T. (1993). The revolution in correctional classification. In American Correctional Association (ACA; Ed.), *Classification: A tool for managing today's offenders* (pp. 1-16). Lanham, MD: ACA.

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Better decisions through science. *Scientific American*, *283*, 82-87.

Teranet IA Inc. (1992). *ModelWare user's manual*. Nanaimo, Canada: Author.

Walters, G. D., White, T. W., & Denney, D. (1991). The Lifestyle Criminality Screening Form: Preliminary data. *Criminal Justice and Behavior*, *18*, 406-441.

Ward, A., & Dockerill, J. (1999). The predictive accuracy of The Violent Offender Treatment Program Risk Assessment Scale. *Criminal Justice and Behavior*, *26*, 125-140.

Webster, C. D., & Menzies, R. J. (1993). Supervision in the deinstitutionalized community. In S. Hodgins (Ed.), *Mental disorder and crime* (pp. 22-38). Newbury Park, CA: Sage.

Webster, C. D., & Polvi, N. H. (1995). Challenging assessments of dangerousness and risk. In J. Ziskin (Ed.), *Coping with psychiatric and psychological testimony* (pp. 221-240). Marina del Rey, CA: Law and Psychology Press.